Computational and Functional Analysis of the Virus-Receptor Interface Reveals Host Range Trade-Offs in New World Arenaviruses

Scott A. Kerr,[a] Eleisha L. Jackson,[b] Oana I. Lungu,[a] Austin G. Meyer,[b] Ann Demogines,[a*] Andrew D. Ellington,[a] George Georgiou,[a,c] Claus O. Wilke,[b] Sara L. Sawyer[a,d]

Departments of Molecular Biosciences,[a] Integrative Biology,[b] and Chemical Engineering,[c] The University of Texas at Austin, Austin, Texas, USA; Department of Molecular, Cellular, and Developmental Biology and the BioFrontiers Institute, University of Colorado Boulder, Boulder, Colorado, USA[d]

ABSTRACT

Animal viruses frequently cause zoonotic disease in humans. As these viruses are highly diverse, evaluating the threat that they pose remains a major challenge, and efficient approaches are needed to rapidly predict virus-host compatibility. Here, we develop a combined computational and experimental approach to assess the compatibility of New World arenaviruses, endemic in rodents, with the host TfR1 entry receptors of different potential new host species. Using signatures of positive selection, we identify a small motif on rodent TfR1 that conveys species specificity to the entry of viruses into cells. However, we show that mutations in this region affect the entry of each arenavirus differently. For example, a human single nucleotide polymorphism (SNP) in this region, L212V, makes human TfR1 a weaker receptor for one arenavirus, Machupo virus, but a stronger receptor for two other arenaviruses, Junin and Sabia viruses. Collectively, these findings set the stage for potential evolutionary trade-offs, where natural selection for resistance to one virus may make humans or rodents susceptible to other arenavirus species. Given the complexity of this host-virus interplay, we propose a computational method to predict these interactions, based on homology modeling and computational docking of the virus-receptor protein-protein interaction. We demonstrate the utility of this model for Machupo virus, for which a suitable cocrystal structural template exists. Our model effectively predicts whether the TfR1 receptors of different species will be functional receptors for Machupo virus entry. Approaches such at this could provide a first step toward computationally predicting the "host jumping" potential of a virus into a new host species.

IMPORTANCE

We demonstrate how evolutionary trade-offs may exist in the dynamic evolutionary interplay between viruses and their hosts, where natural selection for resistance to one virus could make humans or rodents susceptible to other virus species. We present an algorithm that predicts which species have cell surface receptors that make them susceptible to Machupo virus, based on computational docking of protein structures. Few molecular models exist for predicting the risk of spillover of a particular animal virus into humans or new animal populations. Our results suggest that a combination of evolutionary analysis, structural modeling, and experimental verification may provide an efficient approach for screening and assessing the potential spillover risks of viruses circulating in animal populations.

V iruses need to enter cells of a host organism in order to make more copies of themselves. Many viruses interact with protein receptors found on the surface of host cells in order to gain entry into those cells. Like all proteins, these host receptors can vary in sequence from species to species. Viruses tend to be acutely adapted to use the receptor of one species, with the unintentional consequence of being poorly adapted to the receptor encoded by related species. For instance, human but not monkey CD4 serves as a functional receptor for circulating strains of HIV-1 (1, 2). A key event in the emergence of new diseases often involves evolution of the viral genome in a way that renders it compatible with the receptor ortholog encoded by a new host species (reviewed in references 3 to 8). The identification of host genes that impact viral replication in a species-specific fashion is the foundation for understanding why viruses infect the species that they do and essential for understanding the genetic changes that viruses must acquire to infiltrate new species.

The New World arenaviruses, which infect Central and South American rodent species, present an ongoing public health threat. Five different viruses in this family are zoonotic, meaning that they are transmitted from their rodent host species to humans (Fig. 1) (9, 10). Infection in humans can lead to hemorrhagic

fever, and individual outbreaks can have lethality rates as high as 30% (11). The New World arenavirus phylogeny has four major clades, A, B, A/B recombinant, and C, with all zoonotic arenaviruses residing in clade B (12). Transmission to humans likely occurs through direct contact with rodents and through inhalation of aerosolized virions excreted in rodent feces and urine (13). Currently, the geographic ranges of the rodent species that carry these viruses are constrained by specific habitat requirements. However, should arenaviruses ever spread to common species

FIG 1 New World arenaviruses and the species that they infect. A phylogeny of clade B New World arenaviruses is shown (12). The rodent species that are known to be endemically infected with each virus are listed to the right, along with an indication of the country where the rodent/virus pair is found. Asterisks identify zoonotic viruses known to infect humans.

such as the house mouse (*Mus musculus*) or the brown rat (*Rattus norvegicus*), they could become a global threat of immense proportions.

The clade B arenaviruses enter host cells via the host-encoded transferrin receptor 1 (TfR1) (14). The primary function of TfR1 is to mediate the uptake of iron-loaded transferrin, and to a lesser extent H-ferritin (15), from the serum (reviewed in reference 16). The interaction of arenaviruses with TfR1 is mediated by the virus surface spike, or glycoprotein (GP). The viral spike consists of a trimer of GP1/GP2 dimers, and interaction with TfR1 involves GP1 (reviewed in reference 17). A cocrystal structure demonstrating the interaction between human TfR1 and the GP1 of one New World arenavirus, Machupo virus, has been solved (18). Each of the clade B arenaviruses has distinct patterns of compatibility with the TfR1s encoded by various mammalian species (18–22). For instance, Machupo virus enters cells through the TfR1 of its rodent host species, *Calomys callosus*, but not through the TfR1 of the closely related Junin virus host species, *Calomys musculinus* (19). In addition, human TfR1 is a functional receptor only for the five zoonotic New World arenaviruses, not other New World arenaviruses (14, 22, 23). The TfR1s of brown rats and house mice have been found to be nonfunctional entry receptors for all clade B arenaviruses tested (19, 21, 22, 24), in line with the observation that these species have never been found to harbor arenaviruses in nature. Two other families of animal viruses, carnivore parvoviruses and rodent mouse mammary tumor virus (MMTV)-like retroviruses, also use TfR1 to enter cells (25, 26). For all of the virus families that use TfR1, existing evidence suggests that the ability to enter cells through the TfR1 ortholog of a particular species is a necessary criterion for infection in the wild and that viral adaptation is often required to utilize the TfR1 of the new species (7, 27–29).

Recently, we have demonstrated that much of the sequence variability in TfR1 between different rodent and carnivore species is the result of host-virus evolutionary arms races (28, 29). Specifically, *TFR1* has experienced repeated rounds of natural selection for mutations that reduce its functionality as a virus entry receptor. In turn, the viruses that use TfR1 have been counterselected in the gene encoding their surface spike for mutations that reestablish cellular entry through new receptor variants (7, 27). In arms races such as this, codon positions where mutations most potently alter the virus-receptor physical interaction, without affecting protein folding or function, are subject to the strongest positive selection and thus evolve rapidly (30, 31). Arms race dynamics can

be detected by identifying codons that have a significantly higher number of nonsynonymous substitutions than would be expected under a model of neutral or purifying selection (30–32). We previously performed positive selection analysis on an alignment of *TFR1* sequences from various rodent species and found three codons (corresponding to human TfR1 residue positions 205, 209, and 215) near the Machupo virus-binding region that have undergone recurrent positive selection during the speciation of rodents and are therefore evolving at an unusually high rate (28). However, the effect of mutations at these sites in TfR1 has not been tested. Recently, positive selection analysis has proven extraordinarily accurate in identifying the species-specific motifs in entry receptors that govern interactions with viruses (28, 29, 33, 34), but it remains to be explored whether positive selection analysis could add value to our current understanding of the arenavirus-binding surface of TfR1 and the species specificity of this interaction.

Here, we show that TfR1 from the common brown rat functions as a receptor for Machupo virus if residues from the Machupo virus host species, *Calomys callosus*, are substituted into the small protein motif identified by positive selection analysis. We also show that the specific TfR1 residues that dictate Machupo virus entry are not the same residues that govern entry by Chapare and Sabia arenaviruses. Moreover, a human single nucleotide polymorphism (SNP) in this region, L212V, which has previously been shown to make human TfR1 a worse receptor for Machupo virus entry (28), makes it a better receptor for the Junin and Sabia arenaviruses. These findings suggest the potential for evolutionary trade-offs, where selection for resistance for one virus could make humans or rodents susceptible to other arenaviruses in nature. In all of these experiments, we show that the strength of binding affinity between the virus and TfR1 correlates with cellular entry of the virus. Based on this, we develop a structural model for assessing the relative energy of binding of Machupo virus GP1 to TfR1 variants. This model is based on homology modeling of TfR1 variants and subsequent rigid-body docking to Machupo virus GP1. Binding energies between modeled TfR1 variants and Machupo virus GP1 are calculated using the Rosetta interface scoring function (35, 36). We find that the binding-energy predictions made by our model show good agreement with the functional results obtained in this study and others. Our results suggest that a combination of positive selection analysis, structural modeling, and experimental verification may provide an efficient approach for screening and assessing potential spillover risks of viruses circulating in animal populations.

MATERIALS AND METHODS

Cells and plasmids. Human embryonic kidney 293T cells (ATCC) and canine osteosarcoma D17 cells (ATCC) were maintained in Dulbecco's modified Eagle's medium (Cellgro) supplemented with 10% fetal bovine serum (Gibco), 100 units ml^{-1} penicillin, 100 μg ml^{-1} streptomycin, and 2 mM L-glutamine (Cellgro). Human, *Calomys callosus*, and *Rattus norvegicus* TfR1s with a C-terminal FLAG tag were PCR amplified from pcDNA3.1+ vectors (19) and cloned into the Gateway entry vector pCR8 using the pCR8/GW/TOPO TA cloning kit (Invitrogen). The following primers were used to amplify TfR1 for TA cloning: 5′-TTAATACGACT CACTATAGGG-3′ and 5′-TAGAAGGCACAGTCGAGGC-3′. Gateway LR recombination (Invitrogen) was performed to transfer *TFR1* genes from pCR8 into the entry site in a Gateway-converted LPCX retroviral vector. Site-directed mutagenesis of the Rat *TfR1* plasmid was performed to create the rat-short *TFR1* mutant using the QuikChange site-directed

mutagenesis kit (Stratagene) with the following primers: 5′-GGTGACCA TAAATTCAGGTAATGGCGTATACCTAGTGGAGGCTCCTGAGGGT TATGTGGC and 3′-CCTCAGGAGCCTCCACTAGGTATACGCCATT ACCTGAATTTATGGTCACCAAGTTTTGAG. Site-directed mutagenesis was performed to create the rat-long *TfR1* mutant using primers 5′-GGT GACCATAAATGCAAGTAATGGCGTATACCTATTGGAGAGTC CTGAGGGTTATGTGGC and 3′-CCTCAGGACTCTCCAATAGGTAT ACGCCATTACTTGCATTTATGGTCACCAAGTTTTGAG. Site-directed mutagenesis was also performed on human TfR1 to introduce the L212V mutation (28). Plasmids encoding Machupo virus (Carvallo strain) and Junin virus (MC2 strain) GP have been described previously (14). Sabia and Chapare virus GP sequences (NCBI accession numbers YP_089665 and YP_001816782, respectively) in pCAGGS were a gift from Hyeryun Choe.

Generation of stable cell lines. The LPCX:*TfR1* retroviral vectors described above were packaged into virions in 293T cells by cotransfecting them along with pC-VSV-G (a gift from Hyeryun Choe) and the murine leukemia virus (MLV) *gag-pol* packaging plasmid pCS2-mGP (37) using TransIT-293 (Mirus). Two days later, supernatants containing virions were collected, filtered through 0.22-μm filters, and used to infect D17 cells. One milliliter of D17 cells at a concentration of 50,000 cells/ml was plated in 12-well plates 24 h prior to infection. On the day of infection, medium was removed and cells were infected with 2 ml of medium containing 125 to 1,000 μl of virus and Polybrene at a final concentration of 5 μg/ml. Following infection, cells were spinoculated at 1,200 × g for 90 min at 30°C. Twenty-four hours after infection, medium was removed, and fresh medium containing 1.0 μg/ml puromycin was added to select for transduced cells. All TfR1 receptors tested have a C-terminal FLAG tag that is extracellular when the receptor is at the cell surface (21); expression of TfR1 proteins was detected in live cells by flow cytometry using a FLAG antibody conjugated with allophycocyanin (APC; Abcam).

Entry assays. Pseudoviruses used for entry assays (arenavirus GP-pseudotyped MLV recombinant retroviruses) were packaged in 293T cells. TransIT-293 (Mirus) was used to cotransfect the green fluorescent protein (GFP)-encoding transfer vector pQCXIX (BD Biosciences) along with plasmids encoding MLV Gag-Pol and arenavirus GP. After 48 h, supernatants containing viruses were harvested, filtered with 0.22-μm filters, aliquoted, and frozen at −80°C. For entry assays, cell lines stably expressing wild-type and mutant TfR1s were plated at a concentration of 1×10^5 cells per well in 24-well plates and, after 24 h, infected with pseudotyped virus along with Polybrene at a final concentration of 5 μg/ml. The plates were spinoculated with centrifugation at 1,200 × g for 90 min at 30°C. Following spinoculation, cells were washed once with phosphate-buffered saline (PBS) and the medium was replaced. Two days postinfection, cells were trypsinized, treated with 1% paraformaldehyde for 1 h, and labeled using a FLAG antibody conjugated with allophycocyanin (APC; Abcam; catalog no. ab72569) to measure TfR1 expression. Cells were analyzed by flow cytometry. Analysis of flow cytometry data was performed using FlowJo 8.8.6 (TreeStar Inc.). Cells were gated for live populations, and data for 10,000 cells were collected for analysis. For each experiment, cell populations were further gated to represent equivalent mean fluorescence intensity (MFI) for each cell line. This population of cells expressing equivalent amounts of TfR1 was scored for GFP expression (viral entry). Where TfR1 expression is reported, it is reported as the mean fluorescence intensity (MFI) in the analyzed population.

Soluble protein expression. Soluble TfR1s (amino acid positions 117 to 760 in the human TfR1 numbering scheme) were amplified with a 3′ primer that introduces a histidine tag and cloned into pCR4 (Invitrogen); the primers were as follows: 5′-human-GCCCATCTGTCCCGGCCCTG CAGCACGTCGCTTATATTGGGATGACCTGAAGAGAAAG, 5′-rat-G CCCATCTGTCCCGGCCCTGCAGCACGTCGCTTATTTTGGGCAGA CCTCAAAACAC, and 3′-CTATTAATGGTGATGGTGATGGTGATGG TGATGGTGTGATCCACCTCCCTTGTCATCGTCGTCCTTGTAG. This plasmid was used as a template to amplify the TfR1 soluble portion, which was then joined by PCR to a gene block (IDT) containing the

human preprolactin signal sequence (residues 1 to 34), CCGCCACCAT GAACATCAAAGGATCGCCATGGAAAGGGTCCCTCCTGCTGCTGC TGGTGTCAAACCTGCTCCTGTGCCAGAGCGTGGCCCCCTTGCCC ATCTGTCCCGGCCCTGCAGCACGTCGC, using the 5′ primer CCGC CACCATGAACATCAAAGGATCGCCATG and the same 3′ primer as listed above. The product of this reaction was TOPO-TA cloned into pCR8 (Invitrogen) and then Gateway cloned (Invitrogen) into the mammalian expression vector pLPCX. Plasmids were purified via Midi or Maxi Prep purification (Qiagen). For each TfR1, Expi293F cells (Life Technologies) were transfected and incubated for 7 days according to the manufacturer's protocol. A pCDM8 plasmid encoding Machupo virus GP1 fused to the Fc domain of human immunoglobulin (IgG1) protein was received from Sheli Radoshitzky (14). This fusion protein was expressed in Expi293F cells according to the manufacturer's protocol.

Soluble TfR1 and Machupo virus GP1 purification. Supernatant was collected following TfR1 production in Expi293F cells and centrifuged for 20 min at 4,000 relative centrifugal force (RCF) to clear cellular debris. Following centrifugation, the supernatant was passed through an 0.45-μm and then an 0.22-μm filter. An equivalent volume of binding buffer was added (10 mM imidazole, 50 mM NaH_2PO_4, 300 mM NaCl) and incubated with nickel-coated beads for 3 h at 4°C (Complete His-Tag purification resin; Roche). After incubation, the bead-supernatant slurry was passed through 5-ml gravity flow columns (Thermo) two times. The column was washed with 10 ml of wash buffer (20 mM imidazole, 50 mM NaH_2PO_4, 300 mM NaCl). TfR1 was eluted from the column with elution buffer (250 mM imidazole, 50 mM NaH_2PO_4, 300 mM NaCl). Elution fractions containing TfR1 were pooled, concentrated, and buffer exchanged with TfR1 buffer (100 mM NaCl, 100 mM KCl, 5 mM KH_2PO_4, pH 6.7) using Amicon Ultra 100-kDa filter units (Millipore) to capture fully formed TfR1 dimer. Samples were flash-frozen for storage. Purified protein was analyzed by boiling for 10 min in protein loading buffer with 5% beta-mercaptoethanol and resolved on 4 to 20% SDS-PAGE gels (NuSep). The gel was stained with Gel-Code Blue (Life Technologies). For the soluble Machupo virus GP1-Fc, following transfection, the cellular supernatant containing expressed and secreted protein was centrifuged for 20 min at 3,000 RCF at 4°C to clear cellular debris. The supernatant was passed through 0.45-μm filters and then 0.22-μm filters. One milliliter of protein A bead slurry (Thermo) was equilibrated by washing 3 times with 2 ml of PBS, pH 7.4. Filtered supernatant was incubated with equilibrated beads at 4°C on an overhead rotator overnight. Following binding, the bead-supernatant mix was passed through a 5-ml gravity flow column (Thermo). The column was washed with 15 ml wash buffer (PBS, pH 7.4, plus 0.5 M NaCl) and subsequently washed with 15 ml PBS, pH 7.4. Machupo virus GP1-Fc was eluted with elution buffer (PBS, pH 7.4, plus 3.0 MgCl) in 1-ml fractions. Fractions containing protein, as determined by absorbance readings at a 280-nm wavelength via spectrophotometer (NanoDrop; ND-1000), were pooled. The sample was buffer exchanged into PBS (pH 7.4) to reduce the MgCl concentration to 10 mM and flash-frozen for storage. Purified protein was analyzed and stained as described above.

ELISAs. Fifty microliters of TfR1 at 4 μg/ml per well was incubated in PBS overnight at 4°C to bind the protein to the enzyme-linked immunosorbent assay (ELISA) plate. Following binding, wells were washed and blocked with 5% milk in PBS for 2 h at room temperature. Wells were incubated with Machupo virus GP1-Fc in PBS for 1 h at room temperature. Wells were washed and incubated with 1:2,000 horseradish peroxidase (HRP)-conjugated goat antibody raised against human Fc (Jackson ImmunoResearch). Wells were washed, and 50 μl of 1-Step Ultra tetramethylbenzidine (TMB)-ELISA substrate solution (TMB; Thermo Scientific Pierce) was added to the wells and incubated for approximately 15 min. The reaction was quenched with 2 M H_2SO_4, and the plate was read at 450 nm. Curves were fitted using 4-parameter nonlinear regression. For the human transferrin-TfR1 interaction, transferrin-HRP (Jackson ImmunoResearch) in PBS was incubated with 1 mM FeCl(III) for 1 h at room temperature to load

the transferrin with iron. The iron-loaded transferrin-HRP was added to the ELISA wells and incubated at room temperature for 1 hour. Wells were washed, and 50 μl of TMB substrate was added to the wells and incubated for approximately 15 min. The reaction was quenched with 2 M H_2SO_4, and the plate was read at 450 nm. Curves were fitted using 4-parameter nonlinear regression.

Structural homology modeling and docking. We assessed binding affinity in the TfR1-Machupo virus GP1 interface computationally using a strategy of homology modeling and docking. In brief, we first generated homology models of the TfR1-GP1 complex, using the software MODELLER (38, 39) and the published cocrystal structure for human TfR1-Machupo virus GP1 (18). We then refined these models using Rosetta-Dock (36). We modeled the TfR1s of house mouse, brown rat, *Calomys callosus*, and human, as well as various mutants of these receptors: rat-short, rat-long, mouse-human, and human L212V. Our starting point for homology modeling was the crystal structure for the human TfR1-Machupo virus GP1 (PDB code 3KAS) (18), which we cleaned by removing all atoms that did not belong to amino acids. We then made a multiple-sequence alignment using the TfR1s of rat, mouse, human, and *C. callosus* using the software program MAFFT. We gave the aligned sequences of the target TfR1 and the human TfR1 to MODELLER as the input alignment for modeling. Next, we made 100 models, using the basic MODELLER homology modeling protocol. We then used the loop modeling protocol to remodel the loops for each modeled mutant structure. We next took each of these 100 modeled complexes and redocked them in RosettaDock using rigid-body moves (36). The redocking procedure consisted of two steps: prepackaging of the side chains and the actual docking. Modeling parameters were as follows.

For prepackaging, we used the docking prepack protocol in Rosetta with the following flags:

```
-database /path/to/rosetta/database
-l pdb_list.txt #List of structures to
   prepack
-docking:partners A_B
-ex1
-ex2aro, -use_input_sc
-out:file:fullatom
-out:path:pdb ./output_pdbs/
```

For docking, we used the RosettaDock docking protocol with the following flags:

```
-database /path/to/rosetta/database
-l pdb_list.txt #List of prepacked structures
-partners A_B
-dock_pert 3 8
-spin, -ex1
-ex2aro
-use_input_sc
-nstruct 100
-out:file:scorefile
   human_MACV_GP1_docking.fasc #Name of the
   output scorefile
-out:path:pdb ./output_pdbs/ #Path to output
   directory for generated structures
```

Next, we performed a step to refine the docked orientation of the Machupo virus GP1 relative to the new TfR1. For each input structure obtained from MODELLER, we generated 100 docked complexes, for a total of $100 \times 100 = 10,000$ docked complexes for each modeled TfR1. We took the models with the best RosettaDock interface scores to be our representative structures for each modeled complex. Lower interface scores indicate tighter binding affinity. Our complete computational pipeline is available at https://github.com/wilkelab/MACV_TfR1 _modeling.

RESULTS

Signatures of positive selection in *TFR1* refine mapping of the species-specific determinants of Machupo virus entry. A small motif on TfR1 responsible for species-specific interactions with Machupo virus was previously defined by genetic mapping studies. Radoshitzky and colleagues mapped this region using the *TFR1* from house mouse (*Mus musculus*), which does not encode a functional receptor for Machupo virus. They substituted into this gene motifs from the human *TFR1*, which does encode a receptor for Machupo virus, and showed that a 5-amino-acid motif from human TfR1 (positions R208 to L212; blue type in Fig. 2A) conferred a partial increase in Machupo virus entry (19). (For simplicity, all TfR1 coordinates here refer to the human TfR1 numbering.) The cocrystal structure between Machupo virus GP1 and human TfR1 shows that these residues lie predominantly on a beta strand of TfR1 (brown residues in Fig. 2B) in a binding interface where multiple motifs contribute to intermolecular interactions (18). Of the three residue positions that we previously identified to be under positive selection, one (residue 209) falls in this region and two more lie just outside this region (residues 205 and 215, yellow highlighting in Fig. 2A) (28). Based on this, we hypothesized that the species specificity of TfR1 for virus entry might be influenced by TfR1 residues slightly outside the region previously defined.

To test this hypothesis, we examined the TfR1 from the brown rat (*Rattus norvegicus*), which does not function as a receptor for Machupo virus. We focused on the brown rat instead of the house mouse because house mouse TfR1 has a variant amino acid at position 348 that is not shared by most other rodent TfR1s, or by the human TfR1, but which has also been shown to participate in the interaction with Machupo virus (18, 19). We used the TfR1 from the Machupo virus rodent host, *Calomys callosus*, as our permissive control TfR1. We created two chimeric TfR1s in which we swapped residues from the permissive *C. callosus* TfR1 into the nonpermissive rat TfR1 background. In the first chimera, termed rat-short, we replaced five amino acid residues in rat TfR1 (SNIDP) with their respective *C. callosus* residues (NGVYL; red type in Fig. 2A). This clone is structurally similar to the mouse-human TfR1 swap previously described, in that we swapped 5 amino acids including and upstream of the "L" or "P" at position 212. In the second chimera, termed rat-long, we mutated the rat amino acid residues (SGSNIDPVEA) corresponding to human positions 205 to 215 to their respective *C. callosus* TfR1 residues (ASNGVYLLES; red and green type in Fig. 2A). This chimera thus contains the entire stretch of *C. callosus* TfR1 encompassing the three positions that have been targeted by positive selection. (This region is 11 amino acids long in human TfR1 and 10 amino acids long in rodent TfR1.) Stable cell lines expressing each TfR1 were established in canine osteosarcoma cells (D17) so that virus entry assays could be performed. D17 cells were used because the canine TfR1 is not a functional receptor for arenaviruses (19, 21), although modest amounts of background entry are observed for some viruses (reference 22 and demonstrated here). All TfR1 receptors were fused to a C-terminal FLAG tag that is localized extracellularly when the receptor is embedded in the cellular membrane, allowing receptor expression at the cell surface to be detected in live cells by flow cytometry using a FLAG antibody (see Materials and Methods) (Fig. 2C, inset).

Pseudoviruses were used to test arenavirus entry into these
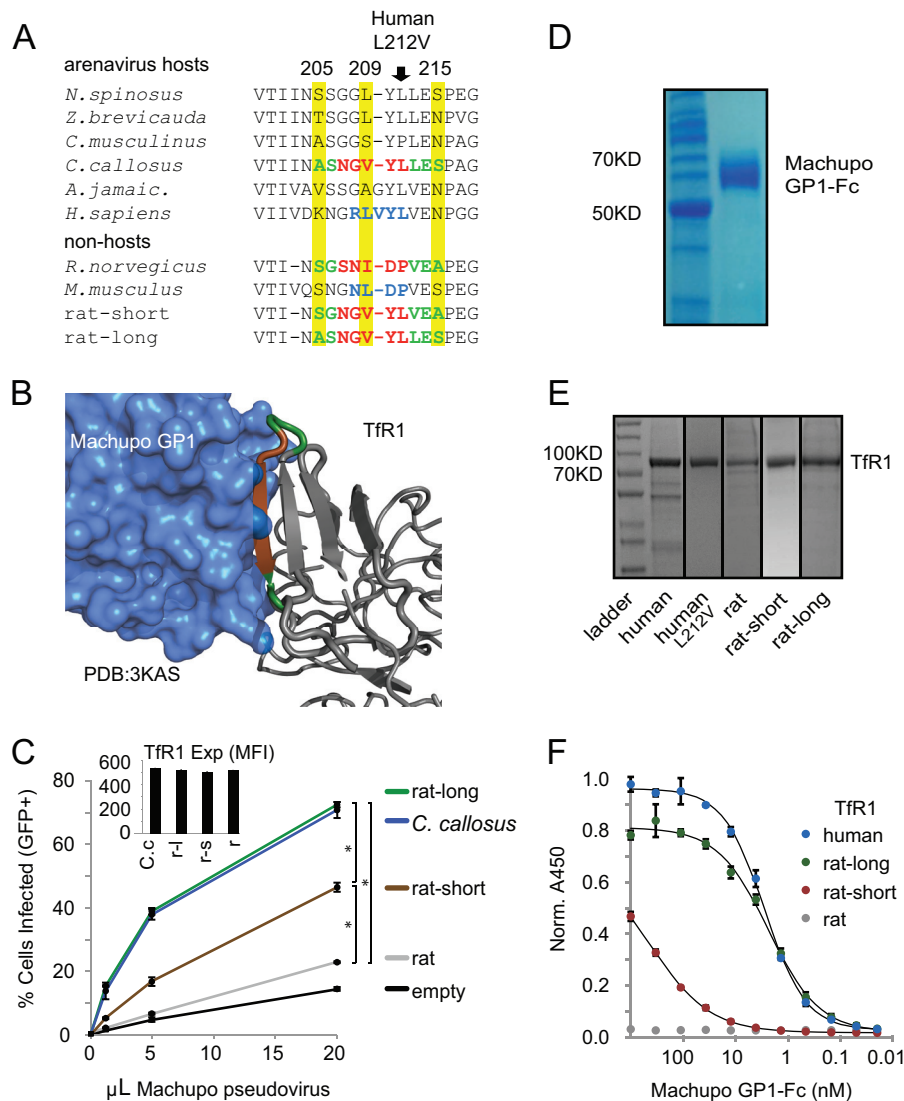
FIG 2 An extended region of TfR1 determines species specificity of Machupo virus entry. (A) An alignment of one portion of the TfR1 apical domain is shown for host and nonhost species. The amino acid numbering corresponds to human TfR1. The three residues experiencing recurrent positive selection, as previously defined (28), are highlighted in yellow. In blue type, a strategy is summarized that was previously used to map the Machupo virus-binding interface on TfR1 (19). In that previous work, this region was swapped from the human TfR1 (RLVYL) into the house mouse (*M. musculus*) TfR1, replacing the corresponding 4 residues (NLDP). In the present work, residues in the *C. callosus* TfR1 were swapped into the rat TfR1 to create the rat-short (red type) and rat-long (red plus green type) chimeric TfR1s shown in the bottom two lines. Also shown is the position of the L212V human SNP discussed in this study. (B) The cocrystal structure shows the interaction between Machupo virus GP1 (blue) and the human TfR1 apical domain (gray) (PDB code 3KAS) (18). The residue positions mutated in the rat-short TfR1 chimera are shown in brown. The residue positions mutated in the rat-long chimera are shown in green and brown. (C) Canine D17 cells were transduced to stably express the wild-type and chimeric TfR1s with C-terminal FLAG tags. Cells were infected with increasing volumes of Machupo pseudovirus (GFP-encoding retrovirus pseudotyped with Machupo virus glycoprotein [GP]). Cells were monitored for GFP expression to determine the percentage of infected cells. Error bars indicate standard deviations from three technical replicates. The experiment was performed 3 times with similar results seen in all experiments; the graph represents data from one experiment. A *t* test was performed to determine if differences between mean values were statistically significant (*, <1e−8). Cell surface expression of TfR1 was monitored with a FLAG antibody via flow cytometry (inset) concurrently with measurement of GFP signal. Expression is given as mean fluorescence intensity (MFI). C.c, *C. callosus*; r-l, rat-long; r-s, rat-short; r, rat. (D) A Coomassie blue-stained SDS-PAGE gel shows purified Machupo virus GP1 fused to the human IgG1 Fc fragment. (E) Coomassie blue-stained SDS-PAGE gels show the five purified TfR1 proteins. (F) ELISA comparing the relative binding affinities of Machupo virus GP1 to each of the purified TfR1s. Purified TfR1 was bound to wells, and then GP1 was incubated at decreasing concentrations to determine the relative binding affinities. Error bars indicate standard deviations from three technical replicates. The experiment was performed 2 times with similar results seen in the two experiments; the graph represents data from one experiment. Curves were fitted using 4-parameter nonlinear regression.

cells. Specifically, Machupo virus glycoprotein (GP) was pseudotyped onto murine leukemia virus (MLV) particles carrying a green fluorescent protein (GFP) gene. Upon virus entry and infection of the cell, the retroviral machinery in the MLV virion drives integration of the GFP gene into the cellular (D17) genome. To assess entry via different versions of TfR1, the D17-TfR1 cell lines were infected with increasing amounts of Machupo pseudovirus and fluorescent cells expressing the GFP reporter were
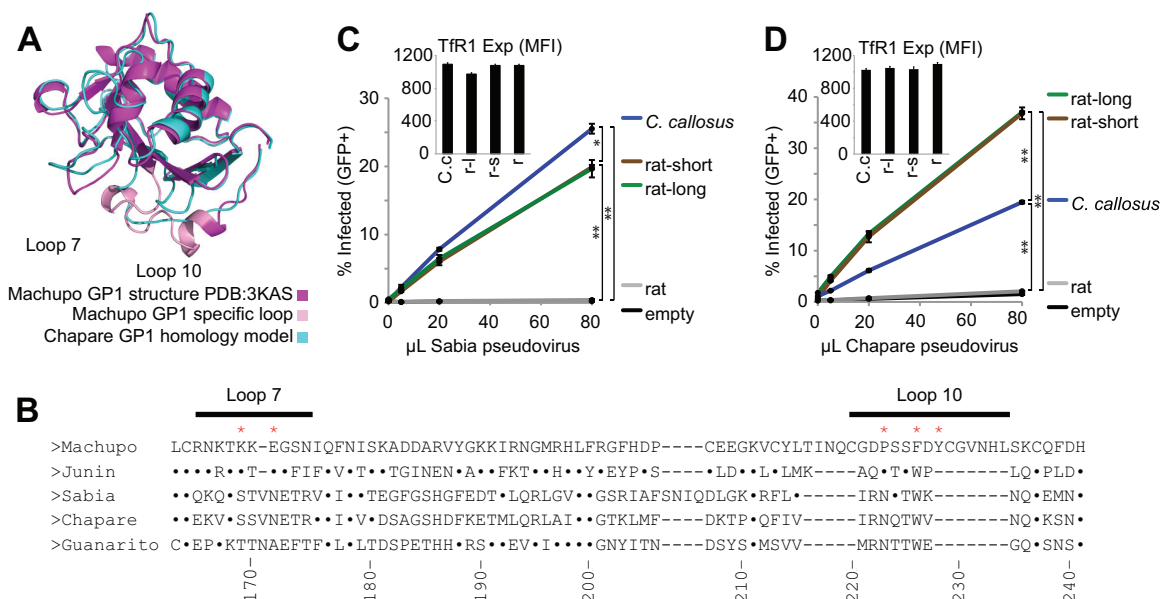
FIG 3 Virus-specific effects on entry through TfR1. (A) The Chapare virus GP1 (blue) was homology modeled on the Machupo virus GP1 crystal structure (magenta; PDB code 3KAS). The Machupo virus GP1 loop 10 and surrounding residues are shown in light pink where divergence in structure is seen in the Chapare virus GP1. Similar results were seen with Sabia and Junin virus models (data not shown). (B) An alignment illustrates the C-terminal region of GP1 from Machupo, Junin, Sabia, Chapare, and Guanarito arenaviruses. Residues identical to those in Machupo virus have been replaced with dots. Machupo virus GP1 sites that make contact with TfR1 (18) are indicated with red asterisks. They fall in loop 7, loop 10, and two upstream regions which are not shown. Loop 10 is longer in Machupo virus than in all other viruses. The numbering scheme is based on the Machupo virus protein. Note that the alignment differs somewhat from a similar alignment shown in reference 18. The differences in these alignments concern the exact placement of insertions and deletions but are within the range to be expected from different alignment approaches. This alignment was generated using MAFFT (47, 48). (C and D) Canine D17 cells were transduced to stably express the wild-type and chimeric TfR1s with C-terminal FLAG tags. Cells were infected with increasing volumes of GFP-encoding retrovirus pseudotyped with Sabia (C) or Chapare (D) virus glycoprotein (GP). Cells were monitored for GFP expression to determine the percentage of infected cells. Error bars indicate standard deviations from three technical replicates. Experiments were performed 2 times, with similar results seen in the two experiments; the graph represents data from one experiment. A t test was performed to determine if differences between mean values were statistically significant (*, $<1e-5$; **, $<1e-10$). Cell surface expression of TfR1 was monitored with a FLAG antibody via flow cytometry (inset) concurrently with measurement of GFP signal. Expression is given as mean fluorescence intensity (MFI). See legend to Fig. 2C for definitions of abbreviations.

counted by flow cytometry 2 days postinfection. As expected, the TfR1 of the natural Machupo virus host species, *C. callosus*, supported high levels of viral entry mediated by the Machupo virus GP (Fig. 2C). As was seen in the mouse-human TfR1 chimera (19), the rat-short TfR1 chimera permitted a moderate amount of entry, greater than that with rat TfR1, confirming that the five amino acid positions introduced from the *C. callosus* TfR1 are partially responsible for Machupo virus GP-mediated entry. The rat-long mutant, whose design was influenced by the evolutionary signatures found in *TFR1*, had an entry phenotype similar to that of the *C. callosus* TfR1, indicating that the species identity of residues in this small 10-amino-acid region more completely defines Machupo virus entry through the TfR1s of different species. However, in some contexts other amino acids, such as that at position 348, may also play a role (18, 19).

To determine whether or not these mutations alter binding affinity between the virus and rat TfR1, we purified a portion of the Machupo virus GP1 (residues 79 to 248) fused to the Fc domain of human IgG1 (Fig. 2D), as was described previously (19). We also purified soluble TfR1 (see Materials and Methods) representing rat as well as the rat-short and rat-long TfR1 variants (Fig. 2E). Despite repeated attempts, we were unable to express and purify the *C. callosus* TfR1. Instead, we purified human TfR1 as a positive control, since Machupo virus can also use this receptor (14). Each TfR1 was applied as a coating onto an ELISA plate, and

purified Machupo virus GP1-Fc was added at decreasing concentrations. Rat TfR1 did not bind GP1 at any concentrations tested (Fig. 2F). The rat-short TfR1 showed moderate binding to GP1 with a 50% effective concentration (EC$_{50}$) of 412 ± 33 nM, while both rat-long TfR1 and human TfR1 bound much more tightly to GP1 with EC$_{50}$s of 2.37 ± 0.18 nM and 2.81 ± 0.12 nM, respectively. Therefore, protein-protein interaction affinities between arenavirus GP and TfR1 correlate with cellular entry, as shown here and previously (18, 19, 21, 22, 40).

Virus-specific effects on cellular entry through TfR1. Residues 205 to 215 in TfR1 interact, in part, with loop 10 in the Machupo virus GP1 (18). As has been previously noted, that loop and surrounding sequence (colored light pink in Fig. 3A) are substantially shorter in length in the GP1s of other arenaviruses (18). Based on this observation, we hypothesized that other New World arenaviruses may have unique contact orientations on TfR1.

To analyze how these structurally different arenavirus GP1s interact with TfR1, we examined our rat-short and rat-long TfR1s for their ability to support entry of another arenavirus. We pseudotyped MLV-based virions with the GP from Sabia virus, a zoonotic virus that is highly diverged from Machupo virus, both phylogenetically (Fig. 1) and in the sequence of loop 10 (Fig. 3B). We again observed that these substitutions into rat TfR1 created a functional receptor for the arenavirus (Fig. 3C). However, in contrast to what we observed with Machupo virus, we found that

rat-short and rat-long TfR1s supported equal levels of entry of Sabia virus. So, residues outside the previously defined region of TfR1 matter for host species compatibility, but to different extents for different arenavirus strains.

Sabia virus GP1 is 4 amino acids longer upstream of loop 10 than other viruses such as Junin, Chapare, and Guanarito viruses (Fig. 3B). To confirm that the unique pattern of entry through these modified rat TfR1s was not a result of this elongated GP1 motif, we tested entry of Chapare virus, another zoonotic virus closely related to Sabia virus (Fig. 1) but lacking the 4-amino-acid insertion (Fig. 3B). We saw the same result, where this virus entered cells equally through the rat-short and rat-long TfR1s (Fig. 3D). Further, we saw additional virus-specific effects in that, for Sabia virus, rat-short and rat-long TfR1s permitted less viral entry than did the *C. callosus* TfR1, while in the case of Chapare virus, they permitted more. This result demonstrates the complexities of predicting host-virus compatibility in nature, where both host and virus genetic variation have functional consequences.

Previously, a human SNP in the arenavirus-binding region of TfR1, L212V (Fig. 4A), was found to reduce entry by Machupo virus (Fig. 4B) (28). It was not previously investigated whether this phenotype was due to reduced interaction affinity. We purified the SNP variant TfR1 L212V protein (Fig. 2E) and again performed an ELISA to measure interaction with Machupo virus GP1. We found that TfR1 L212 bound more tightly to Machupo virus GP1 than did TfR1 L212V (Fig. 4C), in accordance with the entry phenotypes observed. The EC_{50}s of L212 and L212V TfR1 binding to GP1 were 2.1 ± 0.37 nM and >1 μM, respectively. While the interaction between the virus GP1 and TfR1 L212V is very weak, this still appears to result in a modest amount of virus entry (Fig. 4B). This could be because the strength of the interaction is compounded through avidity effects or because the ELISA results underestimate binding due to a high off-rate that may not be as relevant to entry. In contrast, TfR1 binding to iron-loaded human transferrin was not affected by this amino acid change (Fig. 4D).

We next tested two other zoonotic New World arenaviruses for entry through the L212 or L212V allelic variant of human TfR1. We tested one arenavirus that is closely related to Machupo virus (Junin virus) and one that is distantly related (Sabia virus) (Fig. 1). In contrast to what we found for Machupo virus, TfR1 L212V supported higher levels of entry than did the more common TfR1 L212 allelic form, and this was true for both the Junin virus (Fig. 4E) and the Sabia virus (Fig. 4F). In sum, successful pairings between viruses and TfR1 depended on sequence determinants in both the host and virus. For this reason, we next investigated the prospects for computationally screening the effects of mutations via structure-based homology modeling and docking.

Structure-based prediction of host-virus compatibility. It would be useful to be able to computationally predict whether a virus of one species is compatible with the receptor encoded by another species. Toward this goal, we developed a computational pipeline for assessing the effect of species-specific mutations in TfR1 on GP1 binding (Fig. 5A). In brief, we used the Machupo virus GP1-human TfR1 cocrystal structure (PDB code 3KAS) and MODELLER (38, 39) to build homology models of Machupo virus GP1-TfR1 complexes, with species-specific substitutions made in TfR1. The substitutions in TfR1 that we modeled were the same species variants, alleles, and chimeric constructs functionally tested above. We then fine-tuned these models through redocking

with RosettaDock (36). The docking procedure yields characteristic "binding funnels" with the best (i.e., most negative) scores falling into a narrow range of root mean square deviations (RMSDs) relative to the starting conformation (i.e., the models from MODELLER) (Fig. 5A). We used the distribution of the 10 best interface scores as our final measure of the predicted binding strength between the Machupo virus GP1 and the given TfR1.

We tested this method on the eight TfR1 variants discussed in this study: human, human L212V, mouse, the mouse-human chimera, rat, rat-short, rat-long, and *C. callosus* (Fig. 5B). Our computational results showed broad agreement with the experimental findings. TfR1s that functioned as partial or full receptors for Machupo virus in entry assays (colored yellow and green, respectively, in Fig. 5C) had systematically lower (i.e., better) interface scores than did TfR1s that did not function as receptors for Machupo virus (colored red in Fig. 5C). The model accurately predicted the better binding of Machupo virus GP1 to wild-type human TfR1 than to the TfR1 L212V variant. Recapitulating the experimental work done previously (19), it also predicted the successively greater binding strength of Machupo virus GP1 to mouse TfR1, then mouse-human chimeric TfR1, followed by human TfR1. The model predicted that rat-short TfR1 binds to Machupo virus with affinity that is greater than that of rat TfR1 but less than that of *C. callosus* TfR1. One prediction was not completely in line with our experimental observations. Rat-long TfR1 was accurately predicted to bind Machupo virus with an affinity greater than that of rat TfR1 but was not predicted to bind differently than rat-short TfR1, as was our observation. Thus, we found that we could reasonably model Machupo virus GP1 binding to TfR1 in a way that is predictive of whether the virus would be able to enter cells through the TfR1s of different species. We attempted to apply this modeling approach to the other arenavirus strains that we had used in our entry assays, but we found that we were unable to obtain satisfactory homology models and docks due to low template-target sequence identity. Although our approach has this and other limitations, discussed further below, it provides a way to quantify a critical determinant for cross-species transmission of a virus into a new species.

DISCUSSION

Entry receptors are species specific in their interactions with many viruses, and so their characterization in wildlife populations is imperative. Understanding the role of receptors in defining the host range of viruses is important for understanding virus spillover to new hosts and to the identification of model organisms in which to study viral pathogenesis in the lab (2, 41). Here, we propose a combination of computational and experimental approaches to map out the possible host space of a particular virus. Our approach is general and could be applied toward studying the potential spillover risk of many viruses and their interactions with host factors beyond TfR1.

This work has uncovered potential evolutionary trade-offs in the TfR1-arenavirus system. We have shown that a human SNP in TfR1 offers resistance to one virus while making the receptor more vulnerable to other viruses. While this SNP is found in Asia, and not in regions of the world where New World arenaviruses are found (28), one could imagine a situation where selection starts operating on an SNP due to a local infection by a virus to which it offers protection, causing its rise in frequency. This mechanism, however, would create a population that is potentially more vul-
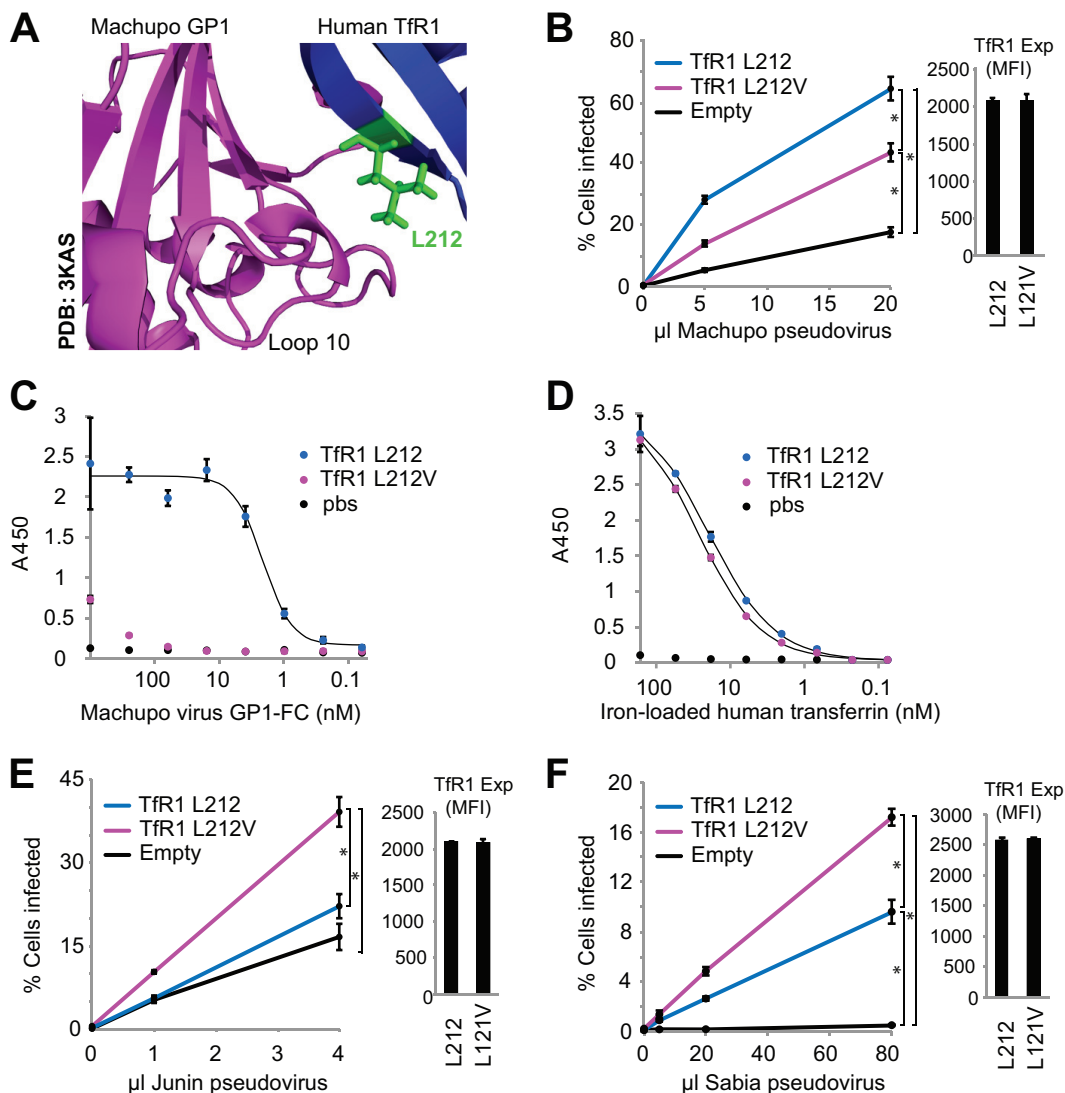
FIG 4 Human TfR1 L212V SNP has an opposite effect on the entry of related arenaviruses. (A) The L212 residue in TfR1 contacts loop 10 in the Machupo virus GP1 (PDB code 3KAS) (18). (B) Canine D17 cells were transduced to stably express the human TfR1 either with L212 or bearing the human SNP mutation, L212V, both with a C-terminal FLAG tag. Cells were infected with increasing volumes of a GFP-encoding retrovirus pseudotyped with Machupo virus glycoprotein. Cells were monitored for GFP expression to determine the percentage of infected cells. Error bars indicate standard deviations from three technical replicates. The experiment was performed 2 times with similar results seen in the two experiments; the graph represents data from one experiment. A *t* test was performed to determine if differences between mean values were statistically significant (*, <1e−3). Cell surface expression of TfR1 was monitored with a FLAG antibody via flow cytometry (right) concurrently with measurement of GFP signal. Expression is given as mean fluorescence intensity (MFI). (C) ELISA comparing the relative binding affinities of the purified human L212 and L212V TfR1s to Machupo virus GP1. Purified TfR1 was bound to wells. GP1 was then incubated at decreasing concentrations to determine the relative binding affinities. Error bars indicate standard deviations from three technical replicates. The experiment was performed 2 times with similar results seen in the two experiments; the graph represents data from one experiment. Curves were fitted using 4-parameter nonlinear regression. pbs, phosphate-buffered saline. (D) ELISA comparing the relative binding affinities of the purified human L212 and L212V TfR1 to iron-loaded human transferrin. For the ELISA, purified TfR1 was bound to wells. Iron-loaded TfR1 was incubated at decreasing concentrations to determine the relative binding affinities. Error bars indicate standard deviations from three technical replicates. The experiment was performed 2 times with similar results seen in the two experiments; the graph represents data from one experiment. Curves were fitted using 4-parameter nonlinear regression. (E and F) The experiment is the same as that in panel B. In this case, cells were infected over increasing volumes with a GFP-encoding retrovirus pseudotyped with Junin (E) or Sabia (F) virus glycoprotein.

nerable to the zoonotic transmission of other viruses. A similar trade-off was evident in the mutational analysis of rat TfR1. Here, we saw that, for Sabia virus, rat-short and rat-long TfR1s were worse receptors than the *C. callosus* TfR1, while in the case of Chapare virus they were better. With receptors, then, a complex fitness landscape can arise where resistance to one pathogen may come at the expense of susceptibility to others.

In the experiments that were performed, a small number of amino acid substitutions in rat TfR1 (rat-short and rat-long) transformed it into a functional receptor for several arenaviruses. This would suggest that SNPs in this region of TfR1, potentially already circulating in rat populations, could lead to individuals that are more susceptible to some arenaviruses and hence might ultimately become carriers of these viruses. This could provide
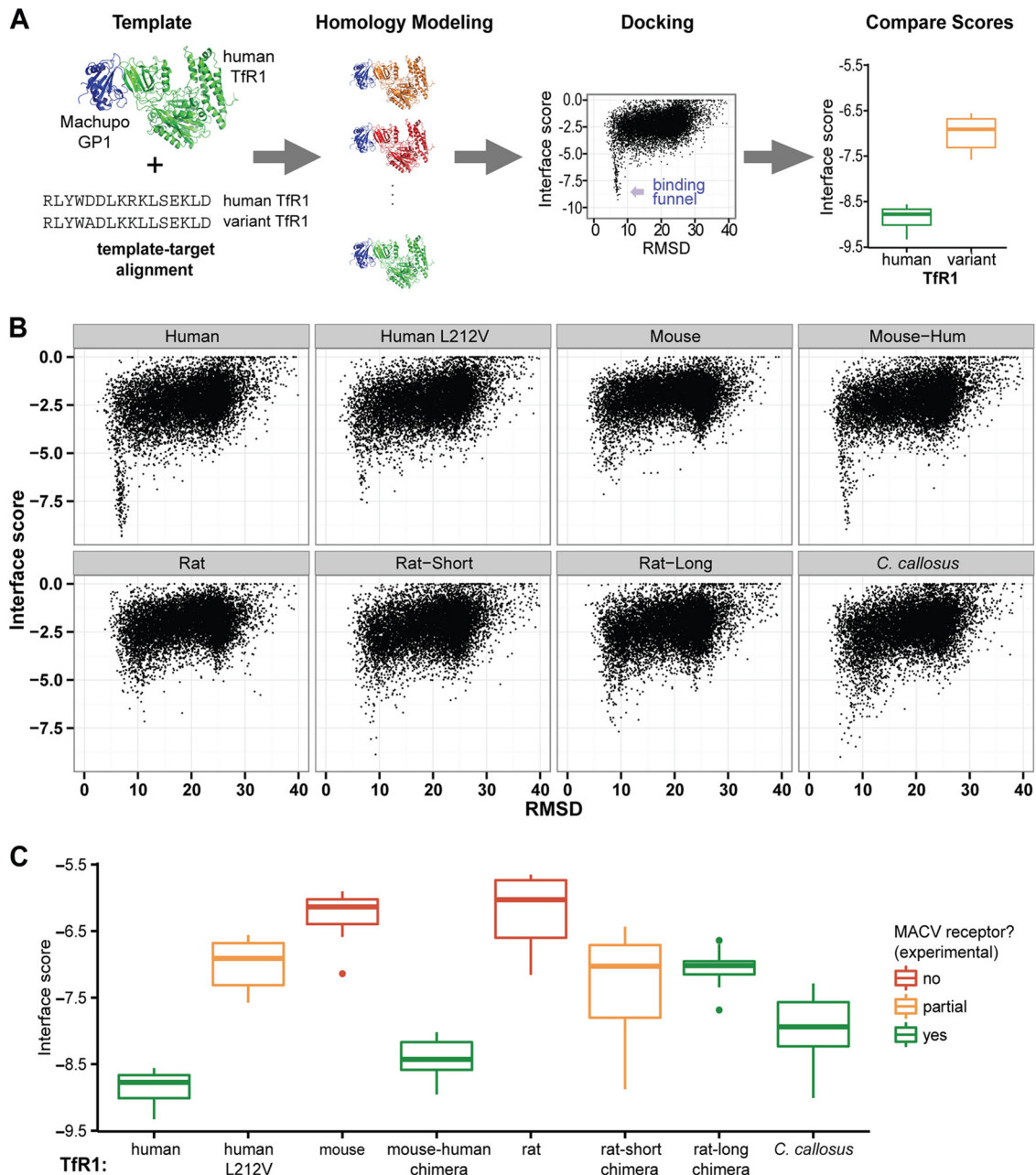
FIG 5 Computational modeling of Machupo virus GP1 binding to different TfR1 variants. (A) The computational pipeline consisted of homology modeling of variant TfR1 to the template cocrystal of the TfR1-GP1 interaction, followed by structural refinement and redocking. The redocking procedure yields a characteristic "binding funnel" of low-energy conformations at around an 8-Å root mean square deviation (RMSD) relative to the starting position before docking. We use the 10 best (i.e., lowest) scores to compare predicted binding strengths, as shown in the graph in the final panel. (B) Binding funnels for the eight models that we analyzed. Each dot shows the Rosetta interface score and the RMSD from the undocked configuration for one of the 10,000 docked models that we generated. (C) Comparison of the 10 best scores across all models considered for each variant TfR1. Each box plot shows the distribution of the interface scores for the 10 best docked conformations for each model. The red, yellow, and green color coding refers to the ability of each TfR1 to act as a receptor for Machupo virus, based on experimental results presented here, or elsewhere for mouse TfR1 (19, 28) and mouse-human TfR1 (19). (One note should be made regarding the color coding of mouse-human TfR1 as green. In the previously published work, the mouse-human TfR1 became fully functional for virus entry only when another mutation, K348N, was also included [19]. On the other hand, K348N alone did not change the functionality of TfR1 for entry [19].)

a conduit for host range expansion into a species that is found globally.

Signatures of positive selection can be a valuable guide to mapping species-specific determinants of virus-receptor compatibility. All that is required to test for positive selection is a sequence alignment of the receptor gene, specifically from species which have served as the long-term viral reservoir. As we have now illustrated with rodent arenaviruses and retroviruses (28), canine parvoviruses (29), bat severe acute respiratory syndrome (SARS)-like coronaviruses (33), and primate lentiviruses (34), the use

of evolutionary analyses to finely map the viral interaction surfaces on entry receptors seems to be accurate in many contexts and can guide the rapid characterization of the host protein receptors that govern species tropism of many emerging viruses. Similar phenomena have also been described for XPR1 gammaretrovirus receptors of rodents and birds (42, 43). Reciprocally, positively selected sites in genes encoding viral glycoproteins can identify the receptor binding site (28, 44).

Indels have accumulated in a region of arenavirus GP1 (loop 10) that contacts TfR1. These changes, combined with positive selection for nonsynonymous mutations in portions of the GP1 that contact TfR1 (28), have resulted in GP1s of arenaviruses being highly divergent. As a consequence, the application of positive selection analysis, phylogenetic studies, and genetic mapping studies to this system becomes difficult, as all of these methods require quality alignments of viral sequences. To help solve this problem, more extensive surveillance of viruses could provide intermediate sequences that resolve the location of indels in viral alignments. Also, multiple cocrystal structures in this system, for instance, the GP1s of several different arenaviruses in complex with the TfR1s of their relative host species, could help elucidate how well or poorly the binding architecture has been preserved.

The structural docking method described here can be used to evaluate the effects of mutations at the receptor-virus interface, similar to what was also done in a recent study with Middle East respiratory syndrome coronavirus (MERS-CoV) and its receptor DPP4 (45). Relative to molecular dynamics methods used previously by us to predict the effect of mutations in the Machupo virus GP1-TfR1 system (46), the docking method described here is much less computationally expensive. While we showed that our structural docking method works well in several contexts in this study, there are limitations. Most obvious is the fact that a cocrystal structure of the interacting proteins in complex is required. While cocrystals are not yet available for many host-virus systems, the number of such structures is quickly growing. It is possible to thread related proteins onto these structures, but we have found that this procedure was more accurate for the host protein, TfR1, than for the viral protein, GP1. Rodent and human TfR1s display only moderate divergence and thus can be reliably homology modeled: all rodent TfR1s analyzed share greater than 70% identity with human TfR1. In contrast, GP1 sequences from different arenaviruses are too divergent to be modeled via homology, due to both the length differences found specifically in the loop 10 binding region and their low amino acid conservation (Machupo virus GP1 residues 87 to 242 share only 25% to 46% identity with the corresponding GP1 residues of Junin, Guanarito, and Sabia viruses) (18). When we tried to model the GP1s of other arenaviruses, they would generally not dock reliably to any TfR1 variant considered (data not shown).

There are several extended applications that can be envisioned for structural docking. First, the TfR1s of rodents with geographic ranges overlapping that of *C. callosus* could be tested to see if there are other rodent species to which spillover of Machupo virus is likely. If these species have broader geographic ranges, this could represent increased risk for humans in those areas. Second, mutations in viral surface glycoproteins could be tested in this model to see which would make the virus better at binding versions of the receptor encoded by new hosts, including humans or new rodent species. These would constitute virus mutations of interest to wildlife surveillance projects. Finally, the results with the L212V

TfR1 SNP suggest that this model can be used to quickly predict the functional effects of human or rodent SNPs. Cocrystal structures describing relevant host-virus protein-protein interactions are the main limitation in this approach. This approach is in its infancy and will be refined by applying it to various problems and learning from its strengths and weaknesses. Investment in the generation of more cocrystal structures demonstrating interactions between host-virus interaction partners would be fruitful, considering the success of the docking approach applied here for Machupo virus GP1 but the poor conservation between arenavirus GP1s.

ACKNOWLEDGMENTS

REFERENCES

1. **Humes D, Emery S, Laws E, Overbaugh J.** 2012. A species-specific amino acid difference in the macaque CD4 receptor restricts replication by global circulating HIV-1 variants representing viruses from recent infection. J Virol **86:**12472–12483. http://dx.doi.org/10.1128/JVI.02176-12.

2. **Meyerson NR, Sharma A, Wilkerson GK, Overbaugh J, Sawyer SL.** 2015. Identification of owl monkey CD4 receptors broadly compatible with early-stage HIV-1 isolates. J Virol **89:**8611–8622. http://dx.doi.org/10.1128/JVI.00890-15.

3. **de Graaf M, Fouchier RAM.** 2014. Role of receptor binding specificity in influenza A virus transmission and pathogenesis. EMBO J **33:**823–841. http://dx.doi.org/10.1002/embj.201387442.

4. **Gerlier D.** 2011. Emerging zoonotic viruses: new lessons on receptor and entry mechanisms. Curr Opin Virol **1:**27–34. http://dx.doi.org/10.1016/j.coviro.2011.05.014.

5. **Graham RL, Baric RS.** 2010. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. J Virol **84:**3134–3146. http://dx.doi.org/10.1128/JVI.01394-09.

6. **Choe H, Jemielity S, Abraham J, Radoshitzky SR, Farzan M.** 2011. Transferrin receptor 1 in the zoonosis and pathogenesis of New World hemorrhagic fever arenaviruses. Curr Opin Microbiol **14:**476–482. http://dx.doi.org/10.1016/j.mib.2011.07.014.

7. **Parrish CR, Kawaoka Y.** 2005. The origins of new pandemic viruses: the acquisition of new host ranges by canine parvovirus and influenza A viruses. Annu Rev Microbiol **59:**553–586. http://dx.doi.org/10.1146/annurev.micro.59.030804.121059.

8. **Parrish CR, Holmes EC, Morens DM, Park E-C, Burke DS, Calisher CH, Laughlin CA, Saif LJ, Daszak P.** 2008. Cross-species virus transmission and the emergence of new epidemic diseases. Microbiol Mol Biol Rev **72:**457–470. http://dx.doi.org/10.1128/MMBR.00004-08.

9. **Delgado S, Erickson BR, Agudo R, Blair PJ, Vallejo E, Albariño CG, Vargas J, Comer JA, Rollin PE, Ksiazek TG, Olson JG, Nichol ST.** 2008. Chapare virus, a newly discovered arenavirus isolated from a fatal hemorrhagic fever case in Bolivia. PLoS Pathog **4:**e1000047. http://dx.doi.org/10.1371/journal.ppat.1000047.

10. **Peters CJ.** 2002. Human infection with arenaviruses in the Americas. Curr Top Microbiol Immunol **262:**65–74.

11. **Borio L, Inglesby T, Peters CJ, Schmaljohn AL, Hughes JM, Jahrling PB, Ksiazek T, Johnson KM, Meyerhoff A, O'Toole T, Ascher MS, Bartlett J, Breman JG, Eitzen EM, Hamburg M, Hauer J, Henderson DA, Johnson RT, Kwik G, Layton M, Lillibridge S, Nabel GJ, Osterholm MT, Perl TM, Russell P, Tonat K, Working Group on Civilian Biodefense.** 2002. Hemorrhagic fever viruses as biological weapons: medical and

public health management. JAMA **287**:2391–2405. http://dx.doi.org/10.1001/jama.287.18.2391.

12. **Charrel RN, Coutard B, Baronti C, Canard B, Nougairede A, Frangeul A, Morin B, Jamal S, Schmidt CL, Hilgenfeld R, Klempa B, de Lamballerie X.** 2011. Arenaviruses and hantaviruses: from epidemiology and genomics to antivirals. Antiviral Res **90**:102–114. http://dx.doi.org/10.1016/j.antiviral.2011.02.009.

13. **LeDuc JW.** 1989. Epidemiology of hemorrhagic fever viruses. Rev Infect Dis **11**(Suppl 4):S730–S735. http://dx.doi.org/10.1093/clinids/11.Supplement_4.S730.

14. **Radoshitzky SR, Abraham J, Spiropoulou CF, Kuhn JH, Nguyen D, Li W, Nagel J, Schmidt PJ, Nunberg JH, Andrews NC, Farzan M, Choe H.** 2007. Transferrin receptor 1 is a cellular receptor for New World haemorrhagic fever arenaviruses. Nature **446**:92–96. http://dx.doi.org/10.1038/nature05539.

15. **Li L, Fang CJ, Ryan JC, Niemi EC, Lebrón JA, Bjorkman PJ, Arase H, Torti FM, Torti SV, Nakamura MC, Seaman WE.** 2010. Binding and uptake of H-ferritin are mediated by human transferrin receptor-1. Proc Natl Acad Sci U S A **107**:3505–3510. http://dx.doi.org/10.1073/pnas.0913192107.

16. **Aisen P.** 2004. Transferrin receptor 1. Int J Biochem Cell Biol **36**:2137–2143. http://dx.doi.org/10.1016/j.biocel.2004.02.007.

17. **Rojek JM, Kunz S.** 2008. Cell entry by human pathogenic arenaviruses. Cell Microbiol **10**:828–835. http://dx.doi.org/10.1111/j.1462-5822.2007.01113.x.

18. **Abraham J, Corbett KD, Farzan M, Choe H, Harrison SC.** 2010. Structural basis for receptor recognition by New World hemorrhagic fever arenaviruses. Nat Struct Mol Biol **17**:438–444. http://dx.doi.org/10.1038/nsmb.1772.

19. **Radoshitzky SR, Kuhn JH, Spiropoulou CF, Albariño CG, Nguyen DP, Salazar-Bravo J, Dorfman T, Lee AS, Wang E, Ross SR, Choe H, Farzan M.** 2008. Receptor determinants of zoonotic transmission of New World hemorrhagic fever arenaviruses. Proc Natl Acad Sci U S A **105**:2664–2669. http://dx.doi.org/10.1073/pnas.0709254105.

20. **Martin VK, Droniou-Bonzom ME, Reignier T, Oldenburg JE, Cox AU, Cannon PM.** 2010. Investigation of clade B New World arenavirus tropism by using chimeric GP1 proteins. J Virol **84**:1176–1182. http://dx.doi.org/10.1128/JVI.01625-09.

21. **Abraham J, Kwong JA, Albariño CG, Lu JG, Radoshitzky SR, Salazar-Bravo J, Farzan M, Spiropoulou CF, Choe H.** 2009. Host-species transferrin receptor 1 orthologs are cellular receptors for nonpathogenic New World clade B arenaviruses. PLoS Pathog **5**:e1000358. http://dx.doi.org/10.1371/journal.ppat.1000358.

22. **Flanagan ML, Oldenburg J, Reignier T, Holt N, Hamilton GA, Martin VK, Cannon PM.** 2008. New world clade B arenaviruses can use transferrin receptor 1 (TfR1)-dependent and -independent entry pathways, and glycoproteins from human pathogenic strains are associated with the use of TfR1. J Virol **82**:938–948. http://dx.doi.org/10.1128/JVI.01397-07.

23. **Helguera G, Jemielity S, Abraham J, Cordo SM, Martinez MG, Rodriguez JA, Bregni C, Wang JJ, Farzan M, Penichet ML, Candurra NA, Choe H.** 2012. An antibody recognizing the apical domain of human transferrin receptor 1 efficiently inhibits the entry of all New World hemorrhagic fever arenaviruses. J Virol **86**:4024–4028. http://dx.doi.org/10.1128/JVI.06397-11.

24. **Cuevas CD, Lavanya M, Wang E, Ross SR.** 2011. Junin virus infects mouse cells and induces innate immune responses. J Virol **85**:11058–11068. http://dx.doi.org/10.1128/JVI.05304-11.

25. **Ross SR, Schofield JJ, Farr CJ, Bucan M.** 2002. Mouse transferrin receptor 1 is the cell entry receptor for mouse mammary tumor virus. Proc Natl Acad Sci U S A **99**:12386–12390. http://dx.doi.org/10.1073/pnas.192360099.

26. **Parker JS, Murphy WJ, Wang D, O'Brien SJ, Parrish CR.** 2001. Canine and feline parvoviruses can use human or feline transferrin receptors to bind, enter, and infect cells. J Virol **75**:3896–3902. http://dx.doi.org/10.1128/JVI.75.8.3896-3902.2001.

27. **Stucker KM, Pagan I, Cifuente JO, Kaelber JT, Lillie TD, Hafenstein S, Holmes EC, Parrish CR.** 2012. The role of evolutionary intermediates in the host adaptation of canine parvovirus. J Virol **86**:1514–1521. http://dx.doi.org/10.1128/JVI.06222-11.

28. **Demogines A, Abraham J, Choe H, Farzan M, Sawyer SL.** 2013. Dual host-virus arms races shape an essential housekeeping protein. PLoS Biol **11**:e1001571. http://dx.doi.org/10.1371/journal.pbio.1001571.

29. **Kaelber JT, Demogines A, Harbison CE, Allison AB, Goodman LB,**

Ortega AN, Sawyer SL, Parrish CR. 2012. Evolutionary reconstructions of the transferrin receptor of caniforms supports canine parvovirus being a re-emerged and not a novel pathogen in dogs. PLoS Pathog **8**:e1002666. http://dx.doi.org/10.1371/journal.ppat.1002666.

30. **Meyerson NR, Sawyer SL.** 2011. Two-stepping through time: mammals and viruses. Trends Microbiol **19**:286–294. http://dx.doi.org/10.1016/j.tim.2011.03.006.

31. **Daugherty MD, Malik HS.** 2012. Rules of engagement: molecular insights from host-virus arms races. Annu Rev Genet **46**:677–700. http://dx.doi.org/10.1146/annurev-genet-110711-155522.

32. **Sironi M, Cagliani R, Forni D, Clerici M.** 2015. Evolutionary insights into host-pathogen interactions from mammalian sequence data. Nat Rev Genet **16**:224–236. http://dx.doi.org/10.1038/nrg3905.

33. **Demogines A, Farzan M, Sawyer SL.** 2012. Evidence for ACE2-utilizing coronaviruses (CoVs) related to severe acute respiratory syndrome CoV in bats. J Virol **86**:6350–6353. http://dx.doi.org/10.1128/JVI.00311-12.

34. **Meyerson NR, Rowley PA, Swan CH, Le DT, Wilkerson GK, Sawyer SL.** 2014. Positive selection of primate genes that promote HIV-1 replication. Virology **454-455**:291–298. http://dx.doi.org/10.1016/j.virol.2014.02.029.

35. **Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, Kaufman K, Renfrew PD, Smith CA, Sheffler W, Davis I, Cooper S, Treuille A, Mandell DJ, Richter F, Ban YE, Fleishman SJ, Corn JE, Kim DE, Lyskov S, Berrondo M, Mentzer S, Popovic Z, Havranek JJ, Karanicolas J, Das R, Meiler J, Kortemme T, Gray JJ, Kuhlman B, Baker D, Bradley P.** 2011. Rosetta3: an object-oriented software suite for the simulation and design of macromolecules. Methods Enzymol **487**:545–574. http://dx.doi.org/10.1016/B978-0-12-381270-4.00019-6.

36. **Chaudhury S, Berrondo M, Weitzner BD, Muthu P, Bergman H, Gray JJ.** 2011. Benchmarking and analysis of protein docking performance in Rosetta v3.2. PLoS One **6**:e22477. http://dx.doi.org/10.1371/journal.pone.0022477.

37. **Yamashita M, Emerman M.** 2004. Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. J Virol **78**:5670–5678. http://dx.doi.org/10.1128/JVI.78.11.5670-5678.2004.

38. **Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-Y, Pieper U, Sali A.** 2006. Comparative protein structure modeling using Modeller. Curr Protoc Bioinformatics Chapter 5:Unit 5.6. http://dx.doi.org/10.1002/0471250953.bi0506s15.

39. **Webb B, Sali A.** 2014. Comparative protein structure modeling using MODELLER. Curr Protoc Bioinformatics **47**:5.6.1–5.6.32. http://dx.doi.org/10.1002/0471250953.bi0506s47.

40. **Radoshitzky SR, Longobardi LE, Kuhn JH, Retterer C, Dong L, Clester JC, Kota K, Carra J, Bavari S.** 2011. Machupo virus glycoprotein determinants for human transferrin receptor 1 binding and cell entry. PLoS One **6**:e21398. http://dx.doi.org/10.1371/journal.pone.0021398.

41. **Sawyer SL, Elde NC.** 2012. A cross-species view on viruses. Curr Opin Virol **2**:561–568. http://dx.doi.org/10.1016/j.coviro.2012.07.003.

42. **Yan Y, Liu Q, Wollenberg K, Martin C, Buckler-White A, Kozak CA.** 2010. Evolution of functional and sequence variants of the mammalian XPR1 receptor for mouse xenotropic gammaretroviruses and the human-derived retrovirus XMRV. J Virol **84**:11970–11980. http://dx.doi.org/10.1128/JVI.01549-10.

43. **Martin C, Buckler-White A, Wollenberg K, Kozak CA.** 2013. The avian XPR1 gammaretrovirus receptor is under positive selection and is disabled in bird species in contact with virus-infected wild mice. J Virol **87**:10094–10104. http://dx.doi.org/10.1128/JVI.01327-13.

44. **Meyer AG, Wilke CO.** 2015. Geometric constraints dominate the antigenic evolution of influenza H3N2 hemagglutinin. PLoS Pathog **11**(5):e1004940. http://dx.doi.org/10.1371/journal.ppat.1004940.

45. **van Doremalen N, Miazgowicz KL, Milne-Price S, Bushmaker T, Robertson S, Scott D, Kinne J, McLellan JS, Zhu J, Munster VJ.** 2014. Host species restriction of Middle East respiratory syndrome coronavirus through its receptor, dipeptidyl peptidase 4. J Virol **88**:9220–9232. http://dx.doi.org/10.1128/JVI.00676-14.

46. **Meyer AG, Sawyer SL, Ellington AD, Wilke CO.** 2014. Analyzing machupo virus-receptor binding by molecular dynamics simulations. PeerJ **2**:e266. http://dx.doi.org/10.7717/peerj.266.

47. **Katoh K, Misawa K, Kuma K-I, Miyata T.** 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res **30**:3059–3066. http://dx.doi.org/10.1093/nar/gkf436.

48. **Katoh K, Kuma K-I, Toh H, Miyata T.** 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res **33**:511–518. http://dx.doi.org/10.1093/nar/gki198.